

Uniform Crime Reporting (UCR) Program Data: A Practitioner's Guide

Jacob Kaplan

2021-03-19

Contents

1	An Overview of the Data	13
1.1	What is a crime?	14
1.1.1	Index crimes	14
1.1.2	Violent crimes	14
1.2	A summary of each UCR dataset	14
1.2.1	Offenses Known and Clearances by Arrest	14
1.2.2	Arrests by Age, Sex, and Race	14
1.2.3	Law Enforcement Officers Killed and Assaulted (LEOKA)	15
1.2.4	Supplementary Homicide Reports (SHR)	15
1.2.5	Hate Crime Data	15
1.2.6	Property Stolen and Recovered (Supplement to Re- turn A)	15
2	Offenses Known and Clearances by Arrest	16
2.1	A brief history of the data	16
2.1.1	Changes in definitions	16
2.2	What does the data look like?	16
2.2.1	Raw data	16
2.2.2	Cleaned data	18

CONTENTS 3

2.3	What variables are in the data?	18
2.3.1	Key variables	18
2.3.2	Known issues with the data	18
2.4	Final thoughts	18
2.5	ORIs - Unique agency identifiers	18
2.6	Hierarchy Rule	21
2.7	Which crimes are included?	21
2.7.1	Index Crimes	21
2.7.2	The problem with using index crimes	24
2.7.3	Rape definition change	24
2.8	Actual offenses, clearances, and unfounded offenses	25
2.8.1	Actual	25
2.8.2	Total Cleared	25
2.8.3	Cleared Where All Offenders Are Under 18	26
2.8.4	Unfounded	26
2.9	Number of months reported	26
3	Arrests by Age, Sex, and Race	27
3.1	A brief history of the data	27
3.1.1	Changes in definitions	27
3.2	What does the data look like?	27
3.2.1	Raw data	27
3.2.2	Cleaned data	27
3.3	What variables are in the data?	31
3.3.1	Key variables	31
3.3.2	Known issues with the data	31
3.4	Final thoughts	31

4	Law Enforcement Officers Killed and Assaulted (LEOKA)	32
4.1	A brief history of the data	33
4.1.1	Changes in definitions	33
4.2	What does the data look like?	33
4.2.1	Raw data	33
4.2.2	Cleaned data	33
4.3	What variables are in the data?	33
4.3.1	Key variables	33
4.3.2	Known issues with the data	34
4.4	Final thoughts	34
5	Supplementary Homicide Reports (SHR)	37
5.1	A brief history of the data	37
5.1.1	Changes in definitions	37
5.2	What does the data look like?	37
5.2.1	Raw data	37
5.2.2	Cleaned data	37
5.3	What variables are in the data?	38
5.3.1	Key variables	38
5.3.2	Known issues with the data	38
5.4	Final thoughts	38
6	Hate Crime Data	41
6.1	A brief history of the data	41
6.1.1	Changes in definitions	41
6.2	What does the data look like?	41
6.2.1	Raw data	41
6.2.2	Cleaned data	41

<i>CONTENTS</i>	5
-----------------	---

6.3	What variables are in the data?	42
6.3.1	Key variables	42
6.3.2	Known issues with the data	42
6.4	Final thoughts	42

7	Property Stolen and Recovered (Supplement to Return A)	45
----------	---	-----------

7.1	A brief history of the data	45
7.1.1	Changes in definitions	45
7.2	What does the data look like?	45
7.2.1	Raw data	45
7.2.2	Cleaned data	45
7.3	What variables are in the data?	48
7.3.1	Key variables	48
7.3.2	Known issues with the data	48
7.4	Final thoughts	48

8	County-Level Detailed Arrest and Offense Data	50
----------	--	-----------

8.1	A brief history of the data	50
8.1.1	Changes in definitions	50
8.2	What does the data look like?	50
8.2.1	Raw data	50
8.2.2	Cleaned data	50
8.3	What variables are in the data?	50
8.3.1	Key variables	50
8.3.2	Known issues with the data	50
8.4	Final thoughts	50

Preface

If you've read an article about crime or arrests in the United States in the last half century, in most cases it was referring to the FBI's Uniform Crime Reporting Program Data, otherwise known as UCR data. UCR data is, with the exception of the more detailed data that only covers murders, a *monthly number of crimes or arrests reported to a single police agency* which is then gathered by the FBI into one file that includes all reporting agencies. Think of your home town. This data will tell you how many crimes were reported for a small number of crimes or how many people (broken down by age, sex, and race) were arrested for a (larger) set of crimes in that city (if the city has multiple police agencies, it will use the agency which is the primary agency on the case, usually the local police department) in a given month. This is a very broad measure of crime, and its uses in research - or understanding crime at all - is fairly limited. Yet it has become - and will likely remain among researchers for at least the next decade - the most important crime data in the United States.

UCR data is important for three reasons:

1. The definitions are standard and all agencies follow them so you can compare across agencies and over time
2. The data is available since 1960 so there is a long period of available data
3. The data is available for most of the 18,000 police agencies in the United States so you can compare across agencies

For most of this book we'll be discussing the caveats of the above reasons - or, more directly, why these assumptions are wrong - but these are the reasons why the data is so influential.

Motivation

By the end of each chapter you should have a firm grasp on the dataset the covered and how to use it properly. However, this book can't possibly cover every potential use case for the data so make sure to carefully examine the data for your own particular use. This benefits you because you'll know your data better and become a better research because of it. This benefits me because it'll increase the quality of research in my field.

I get a lot of emails from people asking questions about this data so part of my motivation for writing this book is to create a single place that answers as many questions as I can about the data. Again, this is among the most commonly used crime datasets and there are still many current papers published with incorrect information about the data (including such simple aspects like what geographic unit data is in and what time unit it is in). So hopefully this book will decrease the amount of misconceptions about this data, increasing overall research quality.¹

Structure of the book

This remainder of this book will be divided into eight chapters: an intro chapter briefly summarizing each dataset and going over overall issues with UCR data, six chapters each covering one of the six UCR datasets, and a final one covering county-level data, a highly flawed but common use of the UCR data. Each chapter will follow the same format: we'll start with a history of the data such as when it first became available and important changes in definitions or variables.

Next we'll discuss what the data looks like initially when you get it from the FBI - literally what it looks like in its fixed-width ASCII format you get from the FBI and what it looks like (what each row and column mean) once it's turned into a useful format that can be read into modern software like R and Stata.² For most of the datasets this is a minor process but for data like

¹Ideally, this will also decrease the number of emails I receive.

²To look at the data in its fixed-width ASCII format I'll use the program Notepad which can open up text files like ASCII files. For looking at the machine-readable format I'll use R since I think it's a bit better looking than viewing it in Stata. In both cases,

the arrest or homicide datasets, the conversion process is harder - and this can actually lead to changes in the resulting data. For example, in an old version of the arrest data that I released, I aggregated certain arrestee ages together since my laptop at the time couldn't handle converting data from ASCII to R and Stata without aggregating the ages (more age groups means more columns which means more computer memory needed). So anyone using that data would have less detailed data than the current dataset.

Understanding how the data moves from its rawest form (which in this case is after cleaning by the FBI) is important for being able to truly understand the data and its caveats. However, this is a fairly technical part of each chapter so feel free to skip it. Next in each chapter, we'll cover the variables included in the data and how to use them properly (including not using them at all) - this will be the bulk of each chapter. We'll end each chapter by briefly summarizing the data, how and when it's useful, and - most importantly - when you shouldn't use it.

Since manuals are boring, I'll try to include graphs and images to try to alleviate the boredom. That said, I don't think it's possible to make it too fun so sorry in advanced. This book is a mix of facts about the data, such as how many years are available, and my opinions about it, such as whether it is reliable. In cases of facts I'll just say a statement - e.g. "the offenses data is available since 1960". In cases of opinion I'll temper the statement by saying something like "in my opinion..."

Citing this book

If this book was useful in your research, please cite it. To cite this book, please use the below citation:

Kaplan J (2021). *Uniform Crime Reporting (UCR) Program Data: A Practitioner's Guide*. <https://github.com/jacobkap/ucrbook>.

BibTeX format:

these will be images included in the chapter - you won't need to follow along or use either program.

Sources of UCR data

There are a few different sources of UCR data available today. First, and probably most commonly used, is the data put together by the [National Archive of Criminal Justice Data \(NACJD\)](#). This a team of out of the University of Michigan who manages a huge number of criminal justice datasets and makes them available to the public. If you have any questions about crime data - UCR or other data - I highly recommend you reach out to them for answers. They have a collection of data and excellent documentation available for UCR data available on their site [here](#). One limitation to their data, however, is that each year of data is available as an individual file meaning that you'll need to concatenate each year together into a single file. Some years also have different column names (generally minor changes like spelling robbery "rob" one year and "robb" the next) which requires more work to standardize before you could concatenate. They also only have data through 2016 which means that the most recent years (UCR data is available through 2019) of data are (as of this writing) unavailable.

Next, and most usable for the general public but limited to researchers, is the FBI's official website [Crime Data Explorer](#). On this site you can chose an agency and see annual crime data (remember, UCR data is monthly so this isn't as detailed as it can be) for certain crimes. This is okay for the general public but only provides a fraction of the data available in the actual data so is really not good for researchers.

Finally, I have my own collection of UCR data [available publicly on openICPSR](#), a site which allows people to submit their data for public access. For each of these datasets I've taken the raw data from the FBI (for early years of homicide data this is actually from NACJD since the FBI's raw data is wrong and can't be parsed. For later years of homicide data this is from the FBI's raw data.) and read it into R. Since the data is only available from the FBI as fixed-width ASCII files, I created a setup file (we'll explain exactly how reading in this kind of data works in the next chapter) and read the data and then very lightly cleaned the data (i.e. only removing extreme outliers like an agency having millions of arsons in a month). For each of these datasets I detail what I've done to the data and briefly summarize the data (i.e. a very short version of this book) on the data's page on openICPSR. The main advantage is that all my data has standard variable names and column names and, for data that is small

enough, provide the data as a single file that has all years. For large datasets like the arrest data I break it down into parts of the data and not all years in a single file. The downside is that I don't provide documentation other than what's on the openICPSR page and only provide data in R and Stata format. I also have a similar site to the FBI's Crime Data Explorer but with more variables available, that site is available [here](#).

It's worth mentioning a final source of UCR information. This is the annual Crimes in the United States report released by the FBI each year around the start of October.³ As an example, here is the [website for the 2019 report](#). In this report is summarized data which in most cases estimates missing data and provides information about national and subnational (though rarely city-level) crime data. As with the FBI's site it is only a fraction of the true data available so is not a very useful source of crime data. Still, this is a very common source of information used by researchers.

Where to find the data used in this book

The data I am using in this book is the cleaned (we'll discuss in more detail exactly what I did to clean each dataset in the dataset's chapter but the short answer is that I did very little.) and concatenated data that I put together from the raw data that the FBI releases. That data is available on my website [here](#). I am hosting this book through GitHub which has a maximum file size allowed that is far smaller than these data so you'll need to go to my site to download the data, it's not available through this book's GitHub repo. For some examples I'm using the data before I cleaned it of outliers (as an example of the outliers present before I removed them) so that data is not publicly available.

NIBRS data

Another source of FBI data, and one sometimes considered part of the UCR data collection, is the National Incident-Based Reporting System (NIBRS)

³They also release a report about the first 6-months of the most recent year of data before the October release but this is generally an estimate from a sample of agencies so is far less useful.

data. Like its name implies this is an incident-level dataset which has detailed information about each incident reported to the police, including incident circumstances, and victim and offender information. This is also the data that the FBI has declared will replace UCR data starting in 2021, meaning that they will no longer collect UCR data and only allow agencies to submit NIBRS data. NIBRS data is a complex and highly rich dataset that deserves its own book to really understand, so I will not be discussing it any further in this book.

About the author

Jacob Kaplan holds a PhD and a master's degree in criminology from the University of Pennsylvania and a bachelor's degree in criminal justice from California State University, Sacramento. His research focuses on Crime Prevention Through Environmental Design (CPTED), specifically on the effect of outdoor lighting on crime. He is the author of several R packages that make it easier to work with data, including [fastDummies](#) and [asciiSetupReader](#). His [website](#) allows easy analysis of crime-related data and he has released over a [dozen crime data sets](#) (primarily FBI UCR data) on openICPSR that he has compiled, cleaned, and made available to the public.

For a list of papers he has written (including working papers), please see [here](#).

For a list of data sets he has cleaned, aggregated, and made public, please see [here](#).

For a list of R packages he has created, please see [here](#).

Chapter 1

An Overview of the Data

One of the first, and most important questions, I think people have about crime is a simple one: is crime going up? Answering it seems simple - you just count up all the crimes that happen in an area and see if that number is bigger than it was in previous times.

However, putting this into practice invites a number of questions, all of which affect how we measure crime. First, we need to define what a crime is? Not philosophically what actions are crimes - or what should be crimes - but literally which of the many thousands of different criminal acts (crimes as defined by state law) should be considered in this measure. Should murder count? Most people would say yes. How about jaywalking or speeding? Many would say probably not. Should marital rape be considered a crime? Now, certainly most people (all, I would hope) would say yes. But in much of the United States it wasn't a crime until the 1970s ([Bennice and Resick, 2003](#); [McMahon-Howard et al., 2009](#)).

Next, we have to know what geographic and time unit to measure crimes at since these decisions determine how precise we can measure crime and when it changed.

The final question is that when a crime occurs, how do we know? That is, when we want to count how many crimes occurred do we ask people how often they've been victimized, do we ask people how often they commit a crime, do we look at crimes reported to police, crimes charged in a criminal court? Each of these measures will likely give different answers as to how

many crimes occurred.¹

The FBI answered all of these questions in 1929 when they began the Uniform Crime Reporting (UCR) Program Data, or UCR data for short. Crime consists of seven crime categories - murder, rape, robbery, aggravated assault, burglary, motor vehicle theft, and theft - that are reported to the police and is collected each month from each agency in the country. These decisions, born primarily out of the resource limitations of 1929 (e.g. no computers), have had a major impact on criminology research. These seven crime categories - known as “Index Crimes” or “Part 1 crimes” (or “Part I” sometimes) - are the ones used to measure crime in many criminology papers, even when the researchers have access to data that covers a broader selection of crimes than these.²

1.1 What is a crime?

1.1.1 Index crimes

1.1.2 Violent crimes

1.2 A summary of each UCR dataset

1.2.1 Offenses Known and Clearances by Arrest

This dataset is the oldest and most widely used of the UCR datasets. It covers

1.2.2 Arrests by Age, Sex, and Race

¹The Bureau of Justice Statistics does measure crime by asking a random sample of people whether they were the victim of a crime. For more on this, please see their National Crime Victimization Survey reports

²Arson is also an index crime but was added after these initial seven were chosen and is not included in the crimes dataset so is generally not included in studies that use index crimes.

1.2.3 Law Enforcement Officers Killed and Assaulted (LEOKA)

1.2.4 Supplementary Homicide Reports (SHR)

1.2.5 Hate Crime Data

1.2.6 Property Stolen and Recovered (Supplement to Return A)

Chapter 2

Offenses Known and Clearances by Arrest

2.1 A brief history of the data

2.1.1 Changes in definitions

2.2 What does the data look like?

2.2.1 Raw data

The data comes as a single fixed-width ASCII file where each row is an agency and the column cover information about the agency (e.g. its name, location) and the monthly number of crimes reported and cleared in that agency. This format makes the data relatively easy to convert into machine-readable formats like R and Stata.

```
knitr::include_graphics('images/offenses_known_raw_ascii_1960.PNG')
```


Figure 2.1: Fixed-width ASCII file for the 1960 Offenses Known and Clearances by Arrest dataset

2.2.2 Cleaned data

```
knitr::include_graphics('images/offenses_known_agency_info_1960.PNG')
```

```
knitr::include_graphics('images/offenses_known_crime_info.PNG')
```

2.3 What variables are in the data?

2.3.1 Key variables

2.3.2 Known issues with the data

2.4 Final thoughts

2.5 ORIs - Unique agency identifiers

In the UCR and other FBI data sets, agencies are identified using **OR**iginating Agency **I**dentifiers or ORIs. These are unique ID codes used to identify an agency. If we used the agency's name we'd end up with some duplicates. For example, if you looked for the Philadelphia Police Department using the agency name, you'd find both the "Philadelphia Police Department" in Pennsylvania and the one in Mississippi.

Each ORI is a 7-digit value starting with the state abbreviation (for some reason the FBI incorrectly puts the abbreviation for Nebraska as NB instead of NE) followed by 5 numbers. In the NIBRS data (another FBI data set) the ORI uses a 9-digit code - expanding the 5 numbers to 7 numbers. When dealing with specific agencies, make sure to use the ORI rather than the agency name to avoid any mistakes.

For an easy way to find the ORI number of an agency, use this [site](#). Type an agency name or an ORI code into the search section and it will return everything that is a match.

ucrbbook - main - RStudio Source Editor

offenses_known_monthly_1960

Filter Cols: 0 - 50

	ori	ori9	agency_name	state	state_abb	year	month	date	number_of_months_missing	last_month_reported	month_missing
1	AK00101	AK0010100	anchorage	alaska	AK	1960	january	1960-01-01	0	december	
2	AK00101	AK0010100	anchorage	alaska	AK	1960	february	1960-02-01	0	december	
3	AK00101	AK0010100	anchorage	alaska	AK	1960	march	1960-03-01	0	december	
4	AK00101	AK0010100	anchorage	alaska	AK	1960	april	1960-04-01	0	december	
5	AK00101	AK0010100	anchorage	alaska	AK	1960	may	1960-05-01	0	december	
6	AK00101	AK0010100	anchorage	alaska	AK	1960	june	1960-06-01	0	december	
7	AK00101	AK0010100	anchorage	alaska	AK	1960	july	1960-07-01	0	december	
8	AK00101	AK0010100	anchorage	alaska	AK	1960	august	1960-08-01	0	december	
9	AK00101	AK0010100	anchorage	alaska	AK	1960	september	1960-09-01	0	december	
10	AK00101	AK0010100	anchorage	alaska	AK	1960	october	1960-10-01	0	december	
11	AK00101	AK0010100	anchorage	alaska	AK	1960	november	1960-11-01	0	december	
12	AK00101	AK0010100	anchorage	alaska	AK	1960	december	1960-12-01	0	december	
13	AK00102	AK0010200	fairbanks	alaska	AK	1960	january	1960-01-01	0	december	
14	AK00102	AK0010200	fairbanks	alaska	AK	1960	february	1960-02-01	0	december	
15	AK00102	AK0010200	fairbanks	alaska	AK	1960	march	1960-03-01	0	december	
16	AK00102	AK0010200	fairbanks	alaska	AK	1960	april	1960-04-01	0	december	
17	AK00102	AK0010200	fairbanks	alaska	AK	1960	may	1960-05-01	0	december	
18	AK00102	AK0010200	fairbanks	alaska	AK	1960	june	1960-06-01	0	december	
19	AK00102	AK0010200	fairbanks	alaska	AK	1960	july	1960-07-01	0	december	
20	AK00102	AK0010200	fairbanks	alaska	AK	1960	august	1960-08-01	0	december	
21	AK00102	AK0010200	fairbanks	alaska	AK	1960	september	1960-09-01	0	december	
22	AK00102	AK0010200	fairbanks	alaska	AK	1960	october	1960-10-01	0	december	
23	AK00102	AK0010200	fairbanks	alaska	AK	1960	november	1960-11-01	0	december	
24	AK00102	AK0010200	fairbanks	alaska	AK	1960	december	1960-12-01	0	december	
25	AK00103	AK0010300	juneau	alaska	AK	1960	january	1960-01-01	0	december	
26	AK00103	AK0010300	juneau	alaska	AK	1960	february	1960-02-01	0	december	
27	AK00103	AK0010300	juneau	alaska	AK	1960	march	1960-03-01	0	december	
28	AK00103	AK0010300	juneau	alaska	AK	1960	april	1960-04-01	0	december	
29	AK00103	AK0010300	juneau	alaska	AK	1960	may	1960-05-01	0	december	
30	AK00103	AK0010300	juneau	alaska	AK	1960	june	1960-06-01	0	december	
31	AK00103	AK0010300	juneau	alaska	AK	1960	july	1960-07-01	0	december	
32	AK00103	AK0010300	juneau	alaska	AK	1960	august	1960-08-01	0	december	
33	AK00103	AK0010300	juneau	alaska	AK	1960	september	1960-09-01	0	december	
34	AK00103	AK0010300	juneau	alaska	AK	1960	october	1960-10-01	0	december	
35	AK00103	AK0010300	juneau	alaska	AK	1960	november	1960-11-01	0	december	
36	AK00103	AK0010300	juneau	alaska	AK	1960	december	1960-12-01	0	december	
37	AK00104	AK0010400	ketchikan	alaska	AK	1960	january	1960-01-01	0	december	

Showing 1 to 38 of 101,424 entries, 175 total columns

Figure 2.2: Agency information for the 1960 Offenses Known and Clearances by Arrest dataset

20CHAPTER 2. OFFENSES KNOWN AND CLEARANCES BY ARREST

[illegible]

Figure 2.3: Crime information for the 1960 Offenses Known and Clearances by Arrest dataset

2.6 Hierarchy Rule

This dataset uses what is called the Hierarchy Rule where only the most serious crime in an incident is reported (except for motor vehicle theft, which is always included). For example if there is an incident where the victim is robbed and then murdered, only the murder is counted as it is considered more serious than the robbery.

How much does this affect our data in practice? Actually relatively little. Though the Hierarchy Rule does mean this data is an under-count, data from other sources indicate that it isn't much of an under count. The FBI's other data set, the National Incident-Based Reporting System (NIBRS) contains every crime that occurs in an incident (i.e. it doesn't use the Hierarchy Rule). Using this we can measure how many crimes the Hierarchy Rule excludes (Most major cities do not report to NIBRS so what we find in NIBRS may not apply to them). In over 90% of incidents, only one crime is committed. Additionally, when people talk about "crime" they usually mean murder which, while incomplete to discuss crime, means the UCR data here is accurate on that measure.

2.7 Which crimes are included?

If you look back at the output when we ran `names(offenses_known_yearly_1960_2017)` you'll see that it produced five broad categories of columns. The first was information about the agency including population and geographic info, then came four columns with the same values except starting with "actual", "tot_clr", "clr_18", and "unfound". Following these starting values were 30 crime categories. We'll discuss what each of those starting values mean in a bit, let's first talk about which crimes are included and what that means for research.

2.7.1 Index Crimes

The Offenses Known and Clearances by Arrest data set contains information on the number of "Index Crimes" (sometimes called Part I crimes) reported to each agency. These index crimes are a collection of eight crimes that, for

22CHAPTER 2. OFFENSES KNOWN AND CLEARANCES BY ARREST

historical reasons based largely by perceived importance in the 1920's when the UCR program was first developed, are used as the primary measure of crime today. Other data sets in the UCR, such as the Arrests by Age, Sex, and Race data and the Hate Crime data have more crimes reported.

The crimes are, in order by the Hierarchy Rule -

1. Homicide

- Murder and non-negligent manslaughter
- Manslaughter by negligence

2. Rape

- Rape
- Attempted rape

3. Robbery

- With a firearm
- With a knife of cutting instrument
- With a dangerous weapon not otherwise specified
- Unarmed - using hands, fists, feet, etc.

4. Aggravated Assault (assault with a weapon or causing serious bodily injury)

- With a firearm
- With a knife of cutting instrument
- With a dangerous weapon not otherwise specified

- Unarmed - using hands, fists, feet, etc.
5. Burglary
 - With forcible entry
 - Without forcible entry
 - Attempted burglary with forcible entry
 6. Theft (other than of a motor vehicle)
 7. Motor Vehicle Theft
 - Cars
 - Trucks and buses
 - Other vehicles
 8. Arson
 9. Simple Assault

For a full definition of each of the index crimes see the FBI's Offense Definitions page [here](#).

Arson is considered an index crime but is not reported in this data - you need to use the separate Arson data set of the UCR to get access to arson counts. The ninth crime on that list, simple assault, is not considered an index crime but is nevertheless included in this data.

Each of the crimes in the list above, and their subcategories, are included in the UCR data. In most reports, however, you'll see them reported as the total number of index crimes, summing up categories 1-7 and reporting that as "crime." These index crimes are often divided into violent index crimes - murder, rape, robbery, and aggravated assault - and property index crimes - burglary, theft, motor vehicle theft.

2.7.2 The problem with using index crimes

The biggest problem with index crimes is that it is simply the sum of 8 (or 7 since arson data usually isn't available) crimes. Index crimes have a huge range in their seriousness - it includes both murder and theft. This is clearly wrong as 100 murders is more serious than 100 thefts. This is especially a problem as less serious crimes (theft mostly) are far more common than more serious crimes (in 2017 there were 1.25 million violent index crimes in the United States. That same year had 5.5 million thefts.). So index crimes under-count the seriousness of crimes. Looking at total index crimes is, in effect, mostly just looking at theft.

This is especially a problem because it hides trends in violent crimes. San Francisco, as an example, has had a huge increase in index crimes in the last several years. When looking closer, that increase is driven almost entirely by the near doubling of theft since 2011. During the same years, violent crime has stayed fairly steady. So the city isn't getting more dangerous but it appears like it is due to just looking at total index crimes.

Many researchers divide index crimes into violent and nonviolent categories, which helps but is still not entirely sufficient. Take Chicago as an example. It is a city infamous for its large number of murders. But as a fraction of index crimes, Chicago has a rounding error worth of murders. Their 653 murders in 2017 is only 0.5% of total index crimes. For violent index crimes, murder makes up 2.2%. What this means is that changes in murder are very difficult to detect. If Chicago had no murders this year, but a less serious crime (such as theft) increased slightly, we couldn't tell from looking at the number of index crimes.

2.7.3 Rape definition change

Starting in 2013, rape has a new, broader definition in the UCR to include oral and anal penetration (by a body part or object) and to allow men to be victims. The previous definition included only forcible intercourse against a woman. As this revised definition is broader than the original one post-2013, rape data is not comparable to pre-2013 data.

2.8 Actual offenses, clearances, and unfounded offenses

For each crime we have four different categories indicating the number of crimes actually committed, the number cleared, and the number determined to not have occurred.

2.8.1 Actual

This is the number of offenses that occurred, simply a count of the number of crimes that month. For example if 10 people are murdered in a city the number of “actual murders” would be 10.

2.8.2 Total Cleared

A crime is cleared when an offender is arrested or when the case is considered cleared by exceptional means. When a single offender for a crime is arrested, that crime is considered cleared. If multiple people committed a crime, only a single person must be arrested for it to be cleared, and as the UCR data is at the offense level, making multiple arrests for an incident only counts as one incident cleared. So if 10 people committed a murder and all 10 were arrested, it would report one murder cleared not 10. If only one of these people are arrested it would still report one murder cleared - the UCR does not even say how many people commit a crime.

A crime is considered exceptionally cleared if the police can identify the offender, have enough evidence to arrest the offender, know where the offender is, but is unable to arrest them. Some examples of this are the death of the offender or when the victim refuses to cooperate in the case.

Unfortunately this data does not differentiate between clearances by arrest or by exceptional means. For a comprehensive report on how this variable can be exploited to exaggerate clearance rates, see [this report by ProPublica](#) on exceptional clearances with rape cases.

2.8.3 Cleared Where All Offenders Are Under 18

This variable is very similar to Total Cleared except is only for offenses in which **every** offender is younger than age 18.

2.8.4 Unfounded

An unfounded crime is one in which a police investigation has determined that the reported crime did not actually happen. For example if the police are called to a possible burglary but later find out that a burglary did not occur, they would put it down as 1 unfounded burglary. This is based on police investigation rather than the decision of any other party such as a coroner, judge, jury, or prosecutor.

2.9 Number of months reported

UCR data is reported monthly though even agencies that decide to report their data may not do so every month. As we don't want to compare an agency which reports 12 months to one that reports fewer, the variable *number_of_months_reported* is way keep only agencies that report 12 months, or deal with those that report fewer.

From our `table()` output it seems that when agencies do report, they tend to do so for all 12 months of the year. However, this variable is seriously flawed, and its name is quite misleading. In reality this variable is actually just whichever the last month reported was. If an agency reported every month of the year, meaning December is the last month, they would have a value of 12. If the agency **only** reported in December, they would also have a value of 12. While there are ways in the monthly data to measure actual number of months reported, these ways are also flawed. So be cautious about this data and particularly the value of this variable.

Chapter 3

Arrests by Age, Sex, and Race

3.1 A brief history of the data

3.1.1 Changes in definitions

3.2 What does the data look like?

3.2.1 Raw data

```
knitr::include_graphics('images/arrests_raw_ascii_1974.PNG')
```

3.2.2 Cleaned data

```
knitr::include_graphics('images/arrests_agency_info.PNG')
```

```
knitr::include_graphics('images/arrests_crime_info.PNG')
```

Figure 3.1: Fixed-width ASCII file for the 1974 Arrests by Age, Sex, and Race dataset

ucrbk - main - RStudio Source Editor

ucr_arrests_monthly_alcohol_or_proper...

Filter Cols: << 0 - 50 >>

	ori	ori9	number_of_months_reported	num_months_forgery	num_months_fraud	num_months_stolen_prop	num_months_dui	num_months_liquor
1	AK00101	AK0010100	12	9	12	2	12	
2	AK00101	AK0010100	12	9	12	2	12	
3	AK00101	AK0010100	12	9	12	2	12	
4	AK00101	AK0010100	12	9	12	2	12	
5	AK00101	AK0010100	12	9	12	2	12	
6	AK00101	AK0010100	12	9	12	2	12	
7	AK00101	AK0010100	12	9	12	2	12	
8	AK00101	AK0010100	12	9	12	2	12	
9	AK00101	AK0010100	12	9	12	2	12	
10	AK00101	AK0010100	12	9	12	2	12	
11	AK00101	AK0010100	12	9	12	2	12	
12	AK00101	AK0010100	12	9	12	2	12	
13	AK00102	AK0010200	12	3	4	3	12	
14	AK00102	AK0010200	12	3	4	3	12	
15	AK00102	AK0010200	12	3	4	3	12	
16	AK00102	AK0010200	12	3	4	3	12	
17	AK00102	AK0010200	12	3	4	3	12	
18	AK00102	AK0010200	12	3	4	3	12	
19	AK00102	AK0010200	12	3	4	3	12	
20	AK00102	AK0010200	12	3	4	3	12	
21	AK00102	AK0010200	12	3	4	3	12	
22	AK00102	AK0010200	12	3	4	3	12	
23	AK00102	AK0010200	12	3	4	3	12	
24	AK00102	AK0010200	12	3	4	3	12	
25	AK00103	AK0010300	12	0	0	5	12	
26	AK00103	AK0010300	12	0	0	5	12	
27	AK00103	AK0010300	12	0	0	5	12	
28	AK00103	AK0010300	12	0	0	5	12	
29	AK00103	AK0010300	12	0	0	5	12	
30	AK00103	AK0010300	12	0	0	5	12	
31	AK00103	AK0010300	12	0	0	5	12	
32	AK00103	AK0010300	12	0	0	5	12	
33	AK00103	AK0010300	12	0	0	5	12	
34	AK00103	AK0010300	12	0	0	5	12	
35	AK00103	AK0010300	12	0	0	5	12	
36	AK00103	AK0010300	12	0	0	5	12	
37	AK00104	AK0010400	1	1	1	1	1	
38	AK00105	AK0010500	12	1	1	1	12	

Showing 1 to 38 of 100,652 entries, 857 total columns

Figure 3.2: Agency information for the 1974 Arrests by Age, Sex, and Race dataset

Figure 3.3: Crime information for the 1974 Arrests by Age, Sex, and Race dataset

3.3 What variables are in the data?

3.3.1 Key variables

3.3.2 Known issues with the data

3.4 Final thoughts

Chapter 4

Law Enforcement Officers Killed and Assaulted (LEOKA)

The Law Enforcement Officers Killed and Assaulted data, often called just by its acronym LEOKA, has two main purposes.¹ First, it provides counts of employees employed by each agency - broken down by if they are civilian employees or sworn officers, and also broken down by gender. And second it measures how many officers were assaulted or killed (including officers who die accidentally such as in a car crash) in a given month - this is broken down into shift type and type (e.g. alone, with a partner, on foot, in a car, etc.), the offender's weapon, and type of call they are responding to (e.g. robbery, disturbance, traffic stop).

The employee information is at the year-level so you know, for example, how many police officers were employed in a given year at an agency, but don't know any more than that such as how many officers were on patrol, were detectives, were in special units, etc. This dataset is commonly used as a measure of police employees and is a generally reliable - though imperfect as we'll see - measure of how many police are employed by a police agency. The second part of this data, measuring assaults and deaths, is more flawed with missing data issues and data error issues (e.g. more officers killed than employed in an agency). We'll get into these issues in more detail during this chapter.

¹This data is also sometimes called the "Police Employees" dataset.

This data, as well as the privately-run site [Officer Down Memorial Page](#) which covers law enforcement officers who have died, has also been used lately in the context of police using force against people out of fear of being harmed by that person. The discussion revolves around whether police are actually in high danger of being harmed by comparing the rate at which officers die to that of other professions. In general they find that police officers are among the most likely profession to die but are not at the top of this measure.

4.1 A brief history of the data

4.1.1 Changes in definitions

4.2 What does the data look like?

4.2.1 Raw data

```
knitr::include_graphics('images/leoka_raw_ascii_1960.PNG')
```

4.2.2 Cleaned data

```
knitr::include_graphics('images/leoka_agency_info_1960.PNG')
```

```
knitr::include_graphics('images/leoka_crime_info_1960.PNG')
```

4.3 What variables are in the data?

4.3.1 Key variables

[illegible]

Figure 4.1: Fixed-width ASCII file for the 1960 Law Enforcement Officers Killed and Assaulted (LEOKA) dataset

4.3.2 Known issues with the data

4.4 Final thoughts

R ucirbook - main - RStudio Source Editor

leoka_monthly_1960

Filter Cols: << 0 - 50 >>

	ori	agency_name	state	state_abb	number_of_months_reported	year	month	date	ori9	fips_state_code	fips_county_code
1	AK00101	anchorage	alaska	AK	11	1960	january	1960-01-01	AK0010100	02	020
2	AK00101	anchorage	alaska	AK	11	1960	february	1960-02-01	AK0010100	02	020
3	AK00101	anchorage	alaska	AK	11	1960	march	1960-03-01	AK0010100	02	020
4	AK00101	anchorage	alaska	AK	11	1960	april	1960-04-01	AK0010100	02	020
5	AK00101	anchorage	alaska	AK	11	1960	may	1960-05-01	AK0010100	02	020
6	AK00101	anchorage	alaska	AK	11	1960	june	1960-06-01	AK0010100	02	020
7	AK00101	anchorage	alaska	AK	11	1960	july	1960-07-01	AK0010100	02	020
8	AK00101	anchorage	alaska	AK	11	1960	august	1960-08-01	AK0010100	02	020
9	AK00101	anchorage	alaska	AK	11	1960	september	1960-09-01	AK0010100	02	020
10	AK00101	anchorage	alaska	AK	11	1960	october	1960-10-01	AK0010100	02	020
11	AK00101	anchorage	alaska	AK	11	1960	november	1960-11-01	AK0010100	02	020
12	AK00101	anchorage	alaska	AK	11	1960	december	1960-12-01	AK0010100	02	020
13	AK00102	fairbanks	alaska	AK	11	1960	january	1960-01-01	AK0010200	02	090
14	AK00102	fairbanks	alaska	AK	11	1960	february	1960-02-01	AK0010200	02	090
15	AK00102	fairbanks	alaska	AK	11	1960	march	1960-03-01	AK0010200	02	090
16	AK00102	fairbanks	alaska	AK	11	1960	april	1960-04-01	AK0010200	02	090
17	AK00102	fairbanks	alaska	AK	11	1960	may	1960-05-01	AK0010200	02	090
18	AK00102	fairbanks	alaska	AK	11	1960	june	1960-06-01	AK0010200	02	090
19	AK00102	fairbanks	alaska	AK	11	1960	july	1960-07-01	AK0010200	02	090
20	AK00102	fairbanks	alaska	AK	11	1960	august	1960-08-01	AK0010200	02	090
21	AK00102	fairbanks	alaska	AK	11	1960	september	1960-09-01	AK0010200	02	090
22	AK00102	fairbanks	alaska	AK	11	1960	october	1960-10-01	AK0010200	02	090
23	AK00102	fairbanks	alaska	AK	11	1960	november	1960-11-01	AK0010200	02	090
24	AK00102	fairbanks	alaska	AK	11	1960	december	1960-12-01	AK0010200	02	090
25	AK00103	juneau	alaska	AK	11	1960	january	1960-01-01	AK0010300	02	110
26	AK00103	juneau	alaska	AK	11	1960	february	1960-02-01	AK0010300	02	110
27	AK00103	juneau	alaska	AK	11	1960	march	1960-03-01	AK0010300	02	110
28	AK00103	juneau	alaska	AK	11	1960	april	1960-04-01	AK0010300	02	110
29	AK00103	juneau	alaska	AK	11	1960	may	1960-05-01	AK0010300	02	110
30	AK00103	juneau	alaska	AK	11	1960	june	1960-06-01	AK0010300	02	110
31	AK00103	juneau	alaska	AK	11	1960	july	1960-07-01	AK0010300	02	110
32	AK00103	juneau	alaska	AK	11	1960	august	1960-08-01	AK0010300	02	110
33	AK00103	juneau	alaska	AK	11	1960	september	1960-09-01	AK0010300	02	110
34	AK00103	juneau	alaska	AK	11	1960	october	1960-10-01	AK0010300	02	110
35	AK00103	juneau	alaska	AK	11	1960	november	1960-11-01	AK0010300	02	110
36	AK00103	juneau	alaska	AK	11	1960	december	1960-12-01	AK0010300	02	110
37	AK00104	ketchikan	alaska	AK	11	1960	january	1960-01-01	AK0010400	02	130
38	AK00104	ketchikan	alaska	AK	11	1960	february	1960-02-01	AK0010400	02	130

Showing 1 to 38 of 101,148 entries, 249 total columns

Figure 4.2: Agency information for the 1960 Law Enforcement Officers Killed and Assaulted (LEOKA) dataset

Chapter 5

Supplementary Homicide Reports (SHR)

5.1 A brief history of the data

5.1.1 Changes in definitions

5.2 What does the data look like?

5.2.1 Raw data

```
knitr::include_graphics('images/shr_raw_ascii_1985.PNG')
```

5.2.2 Cleaned data

```
knitr::include_graphics('images/shr_agency_info.PNG')
```

```

601AL001009A6850001749100370981JEFFERSON      ALA
020517850A001A27MWN31MWN12ST81E000000
601AL001009A6850001749100370981JEFFERSON      ALA
030517850A001A61MWN82MBN12AQ42 000000
601AL001009A6850001749100370981JEFFERSON      ALA
040906850A001A29FWN29MWN13WI60 000000
601AL001009A6850001749100370981JEFFERSON      ALA
050906850A001A45MWN61MWN12AQ60 000000
601AL001009A6850001749100370981JEFFERSON      ALA
050906850A002A03MWN20MWN40SS60 000000
601AL001009A6850001749100370981JEFFERSON      ALA
050906850A003A33MBN33MBN20IL45 000000
601AL001009A6850001749100370981JEFFERSON      ALA
050906850A004A19FBN27FBN20FR45 000000
601AL001009A6850001749100370981JEFFERSON      ALA
071121850A001D30MWN23MWN11AQ18 00100031MWN
601AL001009A6850001749100370981JEFFERSON      ALA
071121850A002A80MBN31FBN11AQ05 000000
601AL001009A6850001749100370981JEFFERSON      ALA
071121850A003A30MBN29MBN20AQ45 000000
601AL001009A6850001749100370981JEFFERSON      ALA

```

Figure 5.1: Fixed-width ASCII file for the 1985 Supplementary Homicide Reports (SHR) dataset

```
knitr::include_graphics('images/shr_crime_info.PNG')
```

5.3 What variables are in the data?

5.3.1 Key variables

5.3.2 Known issues with the data

5.4 Final thoughts

R ucirbook - main - RStudio Source Editor

shr_1976_2018

Filter Cols: << 0 - 50 >>

	ori	ori9	year	state	state_abb	fips_state_code	fips_county_code	fips_state_county_code	fips_place_code	agency_type	cro
1	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
2	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
3	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
4	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
5	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
6	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
7	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
8	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
9	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
10	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
11	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
12	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
13	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
14	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
15	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
16	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
17	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
18	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
19	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
20	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
21	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
22	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
23	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
24	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
25	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
26	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
27	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
28	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
29	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
30	AK00101	AK0010100	2018	alaska	AK	02	020	02020	03000	local police department	an
31	AK00102	AK0010200	2018	alaska	AK	02	090	02090	24230	local police department	fa
32	AK00102	AK0010200	2018	alaska	AK	02	090	02090	24230	local police department	fa
33	AK00102	AK0010200	2018	alaska	AK	02	090	02090	24230	local police department	fa
34	AK00102	AK0010200	2018	alaska	AK	02	090	02090	24230	local police department	fa
35	AK00103	AK0010300	2018	alaska	AK	02	110	02110	36400	local police department	ju
36	AK00103	AK0010300	2018	alaska	AK	02	110	02110	36400	local police department	ju
37	AK00104	AK0010400	2018	alaska	AK	02	130	02130	38970	local police department	ke
38	AK00105	AK0010500	2018	alaska	AK	02	140	02140	40000	local police department	ke

Showing 1 to 333 of 721,956 entries, 159 total columns

Figure 5.2: Agency information for the 1985 Supplementary Homicide Reports (SHR) dataset

ucrbbook - main - RStudio Source Editor

shr_1976_2018

Filter Cols: 0 - 50

incident_number	situation	additional_victim_count	additional_offender_count	victim_1_age	victim_2
1	single victim/single offender	0	0	06	NA
2	single victim/unknown offender(s)	0	0	37	NA
1	single victim/single offender	0	0	33	NA
4	single victim/single offender	0	0	61	NA
3	single victim/single offender	0	0	19	NA
2	single victim/single offender	0	0	19	NA
1	single victim/single offender	0	0	25	NA
3	single victim/single offender	0	0	40	NA
2	single victim/unknown offender(s)	0	0	59	NA
1	single victim/single offender	0	0	45	NA
5	single victim/unknown offender(s)	0	0	23	NA
4	single victim/multiple offenders	0	1	27	NA
1	single victim/single offender	0	0	42	NA
2	single victim/single offender	0	0	20	NA
3	single victim/unknown offender(s)	0	0	52	NA
2	single victim/multiple offenders	0	1	19	NA
1	single victim/unknown offender(s)	0	0	39	NA
2	single victim/single offender	0	0	22	NA
3	single victim/single offender	0	0	48	NA
1	single victim/single offender	0	0	48	NA
2	single victim/single offender	0	0	15	NA
3	single victim/single offender	0	0	31	NA
1	single victim/single offender	0	0	29	NA
2	single victim/multiple offenders	0	1	35	NA
3	single victim/single offender	0	0	24	NA
1	single victim/single offender	0	0	24	NA
1	single victim/single offender	0	0	28	NA
1	single victim/single offender	0	0	36	NA
2	single victim/multiple offenders	0	1	21	NA
4	single victim/single offender	0	0	32	NA
1	single victim/single offender	0	0	7 days to 364 days	NA
1	single victim/single offender	0	0	26	NA
1	single victim/single offender	0	0	27	NA
1	single victim/multiple offenders	0	2	25	NA
1	single victim/single offender	0	0	59	NA
1	single victim/single offender	0	0	55	NA
1	single victim/single offender	0	0	31	NA

Showing 1 to 38 of 721,956 entries, 159 total columns

Figure 5.3: Crime information for the 1985 Supplementary Homicide Reports (SHR) dataset

Chapter 6

Hate Crime Data

This dataset covers crimes that are judged by the police to be motivated by hate. More specifically, they are crimes whose motivation - at least in part - is bias towards a certain person or group of people because of characteristics about them such as race, sexual orientation, or religion.

6.1 A brief history of the data

6.1.1 Changes in definitions

6.2 What does the data look like?

6.2.1 Raw data

```
knitr::include_graphics('images/hate_crimes_raw_ascii_1991.PNG')
```

6.2.2 Cleaned data

```

BH01ALAST0000      19910430      MONTGOMERY, AL
AL8D632N      3490      0000000000
00000000000000000000      00000000000000000000
00000000000000000000      00000000000000000000      00000000001991
ALABAMA HIGHWAY PATROL
BH01AL0010000      1991043019910101BIRMINGHAM, AL
AL9A632N      3070010NA
0001424460370980001407600000000000      00000000000000000000
00000000000000000000      00000000000000000000      00000000001991
JEFFERSON      073
BH01AL0010100      1991043019910101BESSEMER, AL
AL4 631N      3070010NA
0000338980370980000334970000000000      00000000000000000000
00000000000000000000      00000000000000000000      00000000001991
BESSEMER      073
BH01AL0010200      1991043019910101BIRMINGHAM, AL
AL1C631Y      3070010NA
0002693130370980002659680000000000      00000000000000000000
00000000000000000000      00000000000000000000      00000000001991
BIRMINGHAM      073
BH01AL0010300      1991043019910101MOUNTAIN BROOK, AL
AL5 631N      3070010NA
0000200470370980000198100000000000      00000000000000000000
00000000000000000000      00000000000000000000      00000000001991

```

Figure 6.1: Fixed-width ASCII file for the 1991 Hate Crime dataset

```
knitr::include_graphics('images/hate_crimes_agency_info.PNG')
```

```
knitr::include_graphics('images/hate_crimes_crime_info.PNG')
```

6.3 What variables are in the data?

6.3.1 Key variables

6.3.2 Known issues with the data

6.4 Final thoughts

ucrbook - main - RStudio Source Editor

ucr_hate_crimes_1991_2018

Filter Cols: << 0 - 50 >>

	ori	ori9	hate_crime_incident_present_flag	state	state_abb	incident_date	month	month_num	day_of_week	year	population
1	AK00101	AK0010100	one or more hate crime incidents present	alaska	AK	2018-07-05	July	7	Thursday	2018	29
2	AK00102	AK0010200	one or more hate crime incidents present	alaska	AK	2018-03-08	March	3	Thursday	2018	3
3	AK00102	AK0010200	one or more hate crime incidents present	alaska	AK	2018-06-04	June	6	Monday	2018	3
4	AK00102	AK0010200	one or more hate crime incidents present	alaska	AK	2018-08-21	August	8	Tuesday	2018	3
5	AK00103	AK0010300	one or more hate crime incidents present	alaska	AK	2018-07-17	July	7	Tuesday	2018	3
6	AK00103	AK0010300	one or more hate crime incidents present	alaska	AK	2018-09-25	September	9	Tuesday	2018	3
7	AK00104	AK0010400	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
8	AK00105	AK0010500	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
9	AK00106	AK0010600	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
10	AK00107	AK0010700	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
11	AK00108	AK0010800	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
12	AK00109	AK0010900	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
13	AK00110	AK0011000	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
14	AK00111	AK0011100	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
15	AK00112	AK0011200	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
16	AK00113	AK0011300	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
17	AK00114	AK0011400	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
18	AK00116	AK0011600	one or more hate crime incidents present	alaska	AK	2018-05-18	May	5	Friday	2018	
19	AK00117	AK0011700	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
20	AK00118	AK0011800	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
21	AK00119	AK0011900	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
22	AK00120	AK0012000	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
23	AK00121	AK0012100	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
24	AK00122	AK0012200	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
25	AK00123	AK0012300	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
26	AK00124	AK0012400	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
27	AK00125	AK0012500	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
28	AK00126	AK0012600	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
29	AK00127	AK0012700	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
30	AK00128	AK0012800	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
31	AK00130	AK0013000	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
32	AK00131	AK0013100	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
33	AK00132	AK0013200	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
34	AK00133	AK0013300	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
35	AK00135	AK0013500	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
36	AK00142	AK0014200	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	
37	AK00143	AK0014300	no hate crime incidents present	NA	NA	NA	NA	NA	NA	2018	

Showing 1 to 38 of 724,461 entries, 170 total columns

Figure 6.2: Agency information for the 1991 Hate Crime dataset

ucrbk - main - RStudio Source Editor

ucr_hate_crimes_1991_2018

Filter Cols: 70 - 50

	bias_motivation_offense_1	bias_motivation_offense_2	bias_motivation_offense_3	bias_motivation_offense_4	bias_motivation_offense_5	bias_motivation_offense_6
1	anti-black	NA	NA	NA	NA	NA
2	anti-white	NA	NA	NA	NA	NA
3	anti-mental disability	NA	NA	NA	NA	NA
4	anti-white	NA	NA	NA	NA	NA
5	anti-protestant	NA	NA	NA	NA	NA
6	anti-catholic	NA	NA	NA	NA	NA
7	NA	NA	NA	NA	NA	NA
8	NA	NA	NA	NA	NA	NA
9	NA	NA	NA	NA	NA	NA
10	NA	NA	NA	NA	NA	NA
11	NA	NA	NA	NA	NA	NA
12	NA	NA	NA	NA	NA	NA
13	NA	NA	NA	NA	NA	NA
14	NA	NA	NA	NA	NA	NA
15	NA	NA	NA	NA	NA	NA
16	NA	NA	NA	NA	NA	NA
17	NA	NA	NA	NA	NA	NA
18	anti-black	NA	NA	NA	NA	NA
19	NA	NA	NA	NA	NA	NA
20	NA	NA	NA	NA	NA	NA
21	NA	NA	NA	NA	NA	NA
22	NA	NA	NA	NA	NA	NA
23	NA	NA	NA	NA	NA	NA
24	NA	NA	NA	NA	NA	NA
25	NA	NA	NA	NA	NA	NA
26	NA	NA	NA	NA	NA	NA
27	NA	NA	NA	NA	NA	NA
28	NA	NA	NA	NA	NA	NA
29	NA	NA	NA	NA	NA	NA
30	NA	NA	NA	NA	NA	NA
31	NA	NA	NA	NA	NA	NA
32	NA	NA	NA	NA	NA	NA
33	NA	NA	NA	NA	NA	NA
34	NA	NA	NA	NA	NA	NA
35	NA	NA	NA	NA	NA	NA
36	NA	NA	NA	NA	NA	NA
37	NA	NA	NA	NA	NA	NA
38	NA	NA	NA	NA	NA	NA

Showing 1 to 38 of 724,461 entries, 170 total columns

Figure 6.3: Crime information for the 1991 Hate Crime dataset

Chapter 7

Property Stolen and Recovered (Supplement to Return A)

7.1 A brief history of the data

7.1.1 Changes in definitions

7.2 What does the data look like?

7.2.1 Raw data

```
knitr::include_graphics('images/property_stolen_raw_ascii_1960.PNG')
```

7.2.2 Cleaned data

```
knitr::include_graphics('images/property_stolen_agency_info.PNG')
```


ucrbbook - main - RStudio Source Editor

ucr_property_stolen_recovered_monthl...

Filter Cols: << 0 - 50 >>

	ori	ori9	number_of_months_reported	report_indicator	state	agency_name	year	month	date	population	population_group
1	AK00101	AK0010100	12	regular	alaska	anchorage	1960	january	1960-01-01	44237	city 25,000 thru 49,999
2	AK00101	AK0010100	12	regular	alaska	anchorage	1960	february	1960-02-01	44237	city 25,000 thru 49,999
3	AK00101	AK0010100	12	regular	alaska	anchorage	1960	march	1960-03-01	44237	city 25,000 thru 49,999
4	AK00101	AK0010100	12	regular	alaska	anchorage	1960	april	1960-04-01	44237	city 25,000 thru 49,999
5	AK00101	AK0010100	12	regular	alaska	anchorage	1960	may	1960-05-01	44237	city 25,000 thru 49,999
6	AK00101	AK0010100	12	regular	alaska	anchorage	1960	june	1960-06-01	44237	city 25,000 thru 49,999
7	AK00101	AK0010100	12	regular	alaska	anchorage	1960	july	1960-07-01	44237	city 25,000 thru 49,999
8	AK00101	AK0010100	12	regular	alaska	anchorage	1960	august	1960-08-01	44237	city 25,000 thru 49,999
9	AK00101	AK0010100	12	regular	alaska	anchorage	1960	september	1960-09-01	44237	city 25,000 thru 49,999
10	AK00101	AK0010100	12	regular	alaska	anchorage	1960	october	1960-10-01	44237	city 25,000 thru 49,999
11	AK00101	AK0010100	12	regular	alaska	anchorage	1960	november	1960-11-01	44237	city 25,000 thru 49,999
12	AK00101	AK0010100	12	regular	alaska	anchorage	1960	december	1960-12-01	44237	city 25,000 thru 49,999
13	AK00103	AK0010300	0	not reported	alaska	juneau	1960	january	1960-01-01	6797	city 2,500 thru 9,999
14	AK00103	AK0010300	0	not reported	alaska	juneau	1960	february	1960-02-01	6797	city 2,500 thru 9,999
15	AK00103	AK0010300	0	not reported	alaska	juneau	1960	march	1960-03-01	6797	city 2,500 thru 9,999
16	AK00103	AK0010300	0	not reported	alaska	juneau	1960	april	1960-04-01	6797	city 2,500 thru 9,999
17	AK00103	AK0010300	0	not reported	alaska	juneau	1960	may	1960-05-01	6797	city 2,500 thru 9,999
18	AK00103	AK0010300	0	not reported	alaska	juneau	1960	june	1960-06-01	6797	city 2,500 thru 9,999
19	AK00103	AK0010300	0	not reported	alaska	juneau	1960	july	1960-07-01	6797	city 2,500 thru 9,999
20	AK00103	AK0010300	0	not reported	alaska	juneau	1960	august	1960-08-01	6797	city 2,500 thru 9,999
21	AK00103	AK0010300	0	not reported	alaska	juneau	1960	september	1960-09-01	6797	city 2,500 thru 9,999
22	AK00103	AK0010300	0	not reported	alaska	juneau	1960	october	1960-10-01	6797	city 2,500 thru 9,999
23	AK00103	AK0010300	0	not reported	alaska	juneau	1960	november	1960-11-01	6797	city 2,500 thru 9,999
24	AK00103	AK0010300	0	not reported	alaska	juneau	1960	december	1960-12-01	6797	city 2,500 thru 9,999
25	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	january	1960-01-01	6483	city 2,500 thru 9,999
26	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	february	1960-02-01	6483	city 2,500 thru 9,999
27	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	march	1960-03-01	6483	city 2,500 thru 9,999
28	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	april	1960-04-01	6483	city 2,500 thru 9,999
29	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	may	1960-05-01	6483	city 2,500 thru 9,999
30	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	june	1960-06-01	6483	city 2,500 thru 9,999
31	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	july	1960-07-01	6483	city 2,500 thru 9,999
32	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	august	1960-08-01	6483	city 2,500 thru 9,999
33	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	september	1960-09-01	6483	city 2,500 thru 9,999
34	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	october	1960-10-01	6483	city 2,500 thru 9,999
35	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	november	1960-11-01	6483	city 2,500 thru 9,999
36	AK00104	AK0010400	0	not reported	alaska	ketchikan	1960	december	1960-12-01	6483	city 2,500 thru 9,999
37	AK00105	AK0010500	0	not reported	alaska	kodiak	1960	january	1960-01-01	2628	city 2,500 thru 9,999

Showing 1 to 28 of 100,860 entries, 112 total columns

Figure 7.2: Agency information for the 1960 Property Stolen and Recovered dataset

```
knitr::include_graphics('images/property_stolen_crime_info.PNG')
```

7.3 What variables are in the data?

7.3.1 Key variables

7.3.2 Known issues with the data

7.4 Final thoughts

Chapter 8

County-Level Detailed Arrest and Offense Data

8.1 A brief history of the data

8.1.1 Changes in definitions

8.2 What does the data look like?

8.2.1 Raw data

8.2.2 Cleaned data

8.3 What variables are in the data?

8.3.1 Key variables

8.3.2 Known issues with the data

8.4 Final thoughts

Bibliography

- Bennice, J. A. and Resick, P. A. (2003). Marital rape: History, research, and practice. *Trauma, Violence, & Abuse*, 4(3):228–246.
- McMahon-Howard, J., Clay-Warner, J., and Renzulli, L. (2009). Criminalizing spousal rape: The diffusion of legal reforms. *Sociological Perspectives*, 52(4):505–531.